

# MULTI-NODE RECONSTRUCTIBLE PIPELINE COMPUTER

Publication number: JP63147258

Publication date: 1988-06-20

Inventor: DANIERU EMU NOOZENCHIYATSUKU; MAIKERU JII  
RITSUTOMAN

Applicant: UNIV PRINCETON

Classification:

- international: G06F15/16; G06F7/57; G06F9/38; G06F15/177;  
G06F15/80; G06F15/16; G06F7/48; G06F9/38;  
G06F15/76; (IPC1-7): G06F9/38; G06F15/16

- European: G06F7/57; G06F9/38; G06F15/80B

Application number: JP19870285643 19871113

Priority number(s): US19860931549 19861114

Also published as:

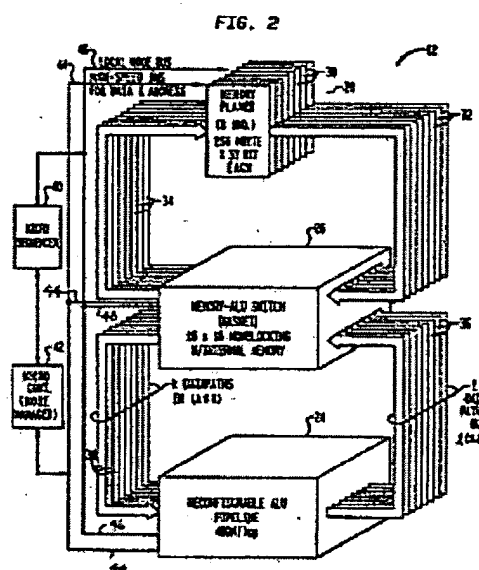
EP0268435 (A2)  
US4811214 (A1)  
EP0268435 (A3)  
EP0268435 (B1)  
ES2070825T (T3)

Report a data error here

Abstract not available for JP63147258

Abstract of corresponding document: **EP0268435**

A multinode parallel-processing computer comprises a plurality of interconnected, large capacity nodes (12) each including a reconfigurable pipeline of functional units such as Integer Arithmetic Logic Processors, Floating Point Arithmetic Processors, Special Purpose Processors, etc.. The reconfigurable pipeline of each node (12) is connected to a multiplane memory (28) by a Memory-Alu Switch NETWORK (MASNET) (26). The reconfigurable pipeline conveniently includes three basic substructures formed from functional units which have been found to be sufficient to perform the bulk of all calculations. The MASNET (26) controls the flow of signals from the memory planes (30) to the reconfigurable pipeline (24) and vice versa. The nodes (12) are connectable together by an internode data router (Hyperspace router) so as to form a hypercube configuration. The capability of the nodes (12) to conditionally reconfigure the pipeline (24) at each tick of the clock, without requiring a pipeline flush, permits many powerful algorithms to be implemented directly.



Data supplied from the esp@cenet database - Worldwide

## ⑫ 公開特許公報(A)

昭63-147258

⑮ Int.Cl.<sup>4</sup> 識別記号 庁内整理番号 ⑬ 公開 昭和63年(1988)6月20日  
 G 06 F 15/16 3 9 0 T-2116-5B  
 9/38 3 7 0 7361-5B  
 15/16 4 0 0 N-2116-5B 審査請求 未請求 発明の数 2 (全15頁)

⑭ 発明の名称 マルチノード再構成可能バイブラインコンピュータ

⑯ 特 願 昭62-285643

⑰ 出 願 昭62(1987)11月13日

優先権主張 ⑱ 1986年11月14日 ⑲ 米国(US) ⑳ 931,549

㉑ 発 明 者 ダニエル・エム・ノー アメリカ合衆国 ニュージャージー州 08619 マーサー  
 ゼンチャツク ビル、レイザーバツク・ドライブ 6

㉒ 発 明 者 マイケル・ジー・リッ アメリカ合衆国 ペンシルバニア州 19107 フィラデル  
 トマン フィア、サウス・ナインス・ストリート 240

㉓ 出 願 人 ザ・トラステイズ・ アメリカ合衆国 ニュージャージー州 08544 プリンス  
 オブ・プリンストン・ トン、ニュー・サウス・ビルディング 3  
 ユニバーシティー

㉔ 代 理 人 弁理士 木下 洋平 外1名

## 明 細 書

1. 発明の名称 マルチノード再構成可能  
 バイブラインコンピュータ

## 2. 特許請求の範囲

(1) それぞれが再構成可能論理演算バイブライン  
 ユニットを含む複数のノードと、

前記ノードの間においてデータを経路指定す  
 るためのルーティング手段を有してなる、

マルチノード並行処理コンピュータ装置。

(2) 前記各ノードがさらに内部メモリーを有して  
 いる、特許請求の範囲第1項に記載の装置。

(3) 各ノードが、さらに前記内部メモリーから前  
 記マスネットを通過して前記再構成可能ALUバイ  
 ブラインユニットに、及び前記再構成可能AL  
 Uバイブラインユニットから前記マスネット  
 を通過して前記内部メモリーにデータを経路指定  
 する為のメモリーALUスイッチネットワーク  
 (マスネット)を有している、特許請求の範囲  
 第2項に記載の装置。

(4) 前記再構成可能ALUバイブラインユニット

が、

入力及び出力を有する第1の構成において永  
 続的に結合されているプログラマブルプロセッ  
 サの第1のグループと、

同様に入力と出力を有し、前記第1の構成と  
 は異なる第2の構成において永続的に接続され  
 ているプログラマブルプロセッサの第2のグル  
 ープと、

前記第1及び第2のグループを互いに選択的  
 に接続するためのALUバイブライン構成スイ  
 ッチングネットワーク手段(フローネット)と、  
 を有し、

かくして、前記再構成可能ALUバイブライ  
 ユニットが前記フローネットからの命令に応  
 じて選択的に異なる演算を行うことを特徴とす  
 る、

特許請求の範囲第3項に記載の装置。

(5) 前記第1のグループのプログラマブルプロセ  
 ッサが、

少なくとも2つの入力と少なくとも1つの出

力を有する第1のプロセッサと、

少なくとも2つの入力と少なくとも1つの出力を有する第2のプロセッサと、

同じく出力を有し前記第1及び第2のプログラマブルプロセッサの出力に永続的に接続されている2つの入力を有する第3のプログラマブルプロセッサを有し、

かくして、前記第1のグループへの入力が前記第1及び第2のプログラマブルプロセッサの入力を構成しかつ前記第1のグループの出力が第3のプログラマブルプロセッサの出力を構成するようになっている、

特許請求の範囲第4項記載の装置。

(6) 前記第2のプログラマブルプロセッサが、

少なくとも2つの入力と少なくとも1つの出力を有する第4のプログラマブルプロセッサと、

2つの入力と1つの出力を有する第5のプログラマブルプロセッサを有し、前記第5のプログラマブルプロセッサの前記入力の1つは前記第4のプログラマブルプロセッサの出力に永続

的に接続されており、かくして、  
ラマブルプロセッサの第2のグループのものに対する比がほぼ1.5乃至2.0:1.0の範囲にある、特許請求の範囲第7項記載の装置。

(9) 前記プログラマブルプロセッサの第2のグループのもの前記プログラマブルプロセッサの第3のグループのものに対する比がほぼ2.0:1.0である、特許請求の範囲第8項記載の装置。

00 前記内部メモリーが複数のメモリープレーンからなる、特許請求の範囲第9項記載の装置。

01 各メモリープレーンが、

メインメモリーバンクと、

前記メインメモリーバンクから、また、に対してデータを伝達するためのアドレスマルチプレクサと、

前記メモリーバンクと前記アドレスマルチプレクサとの間に接続されている先取アドレスレジスターと、

前記アセンブリバンクをアランダムアクセス方式において操作するための前記アドレスマルチプレクサに接続されている翻訳テーブル手

段的に接続されており、かくして、

前記第2のグループの入力が前記第4のプログラマブルプロセッサに対する2つの入力を構成し、且つ前記第5のプログラマブルプロセッサの入力であって前記第4のプログラマブルプロセッサの出力に接続されていないもの、そして、第2のグループの出力が、前記第4のプログラマブルプロセッサの出力を構成するようになっている、

特許請求の範囲第5項記載の装置。

(7) 前記再構成可能ALUパイプラインユニットが更に前記第1及び第2のプログラマブルプロセッサに選択的に接続するために、前記フローネットに接続されるそれぞれのプログラマブルプロセッサからなるプログラマブルプロセッサの第3のグループを有している、

特許請求の範囲第6項記載の装置。

(8) ある与えられた再構成可能ALUパイプラインユニットにおける前記プログラマブルプロセッサの第1のグループのものの前記プログ

段、を有する、

特許請求の範囲第10項記載の装置。

02 各ノードが、さらに、前記内部メモリー、マスネット及び前記再構成可能ALUパイプラインユニットの間におけるデータのブロッキングを支配するための前記内部メモリー、マスネット及び再構成可能のALUパイプラインユニットに接続されているマイクロシーケンサー手段を有する、特許請求の範囲第11項記載の装置。

03 各ノードが、さらに、前記内部メモリー、マスネット、再構成可能ALUパイプラインを初期化し、ベリファイするための前記内部メモリー及びマスネット、前記再構成可能ALUパイプラインに接続されているマイクロコントローラを有する、特許請求の範囲第12項記載の装置。

04 前記マスネットがどのような入力のどのような出力に対する結合関係もノンブロッキングとなるように、ビーンズスイッチングネットワーク

クにおいて結合され、パイプライン化されている複数のレジスタファイルを有する、特許請求の範囲第13項記載の装置。

- (59) BPDデータを特定の目的地ノードに経路指定するために前記ルーティング手段、前記マシネットの間に結合されているバウンダリ・ポイント・デフェニション、キャッシュ手段を具え、かくして前記装置が全体的アドレス、PBDアドレスモードをサポートするようになっている、特許請求の範囲第14項記載の装置。

- (60) 前記ノードにデータと命令を送るための前置コンピュータと前記前置コンピュータに結合しうるオフラインの大容量記憶手段を有する、特許請求の範囲第15項記載の装置。

- (71) 前記ノードが、ブールハイパーキューブのトポロジーにおいて結合され、かつその数が1〜128の間において変化する、特許請求の範囲第16項記載の装置。

- (80) 入力及び出力を有する第1の構成において永続的に結合されているプログラマブルプロセ

ッサを有し、前記第1のグループの入力が前記第1及び第2のプログラマブルプロセッサの入力を構成し、且つ前記第1のグループの出力が前記第3のプログラマブルプロセッサの出力を構成するようになっている、

特許請求の範囲第18項記載の装置。

- (90) 前記プログラマブルプロセッサの第2のグループのものが、

2つの入力及び1つの出力を有する第4のプログラマブルプロセッサと、

2つの入力及び1つの出力を有する第5のプログラマブルプロセッサであって、前記第5のプログラマブルプロセッサの入力の1つが前記第4のプログラマブルプロセッサの出力に永続的に連結されており、

前記第2のグループの入力が前記第4のプログラマブルプロセッサに対する2つの入力を構成し、且つ前記第5のプログラマブルプロセッサに対する入力であって前記第4のプログラマブルプロセッサの出力に連結されているもの、

サの第1のグループと、

同様に入力及び出力を有して前記第1の構成とは異なる、第2の構成において永続的に接続されているプログラマブルプロセッサの第2のグループと、

前記第1及び第2のグループを選択的に接合するためのスイッチング手段（フローネット）を有し、

前記スイッチング手段からの命令に応じて選択的に異なる演算を行うようにした、

再構成可能コンピュータ装置。

- (100) 前記プログラマブルプロセッサの第1のグループのものが、少なくとも2つの入力及び少なくとも1つの出力を有する第1のプログラマブルプロセッサと、

少なくとも2つの入力と1つの出力を有するプログラマブルプロセッサと、

同様に出力を有して前記第1及び第2のプログラマブルプロセッサの出力に永続的に連結されている2つの入力と有する第3のプログラマ

ブルプロセッサの出力が、前記第5のプログラマブルプロセッサの出力を構成している、

特許請求の範囲第19項記載の装置。

- (21) 2つの入力及び1つの出力を有するそれぞれのプログラマブルプロセッサを有してなる第3のプログラマブルプロセッサのグループであって、その第3のグループは前記第1及び第2のグループと選択的に接続するため、前記スイッチング手段に結合されている、特許請求の範囲第20項記載の装置。

- (22) 前記内部メモリから前記マシネットを通過して前記スイッチング手段及び前記スイッチング手段から前記マシネットを通して前記内部メモリにデータを移送するためのメモリ・A・L・Uスイッチネットワーク手段（マシネット）をさらに有する、特許請求の範囲第21項記載の装置。

- (23) 前記内部メモリ、マシネット及びスイッチング手段の間においてデータのブロッキングを

支配するため、前記内部メモリー、マスネット、及びスイッチング手段に連結されているマイクロセンサ手段をさらに有する、特許請求の範囲第22項記載の装置。

(24) 前記内部メモリー、マスネット及びスイッチング手段を初期化し、その状態をベリファイするための前記内部メモリー、マスネット、スイッチング手段に連結されているマイクロコントローラ手段をさらに有する、特許請求の範囲第23項記載の装置。

(25) 前記プロセッサの少なくともいくつかのものが、浮動小数点算術プロセッサからなる、特許請求の範囲第18項記載の装置。

(26) 前記プロセッサの少なくともいくつかのものが、整数算術論理プロセッサである、特許請求の範囲第18項記載の装置。

(27) 前記プログラマブルプロセッサの第1のグループのものの、前記プログラマブルプロセッサの第2のグループのものに対する比がほぼ1.5乃至2.0:1.0の範囲にある、特許請求の範囲

第18項記載の装置。

(28) 前記プログラマブルプロセッサの第2のグループのものの、前記プログラマブルプロセッサの第3のグループのものに対する比がほぼ2.0:1.0である、特許請求の範囲第21項記載の装置。

### 3. 発明の詳細な説明

#### 産業上の利用分野

この発明は、各ノード(node)が、マルチプル、独立メモリープレーンに多機能メモリー—ALUスイッチネットワーク(MASNET)を通じて連結され、マルチプルノードがハイパーキューブ(hypercube)トポロジーにおいて連結されている、再構成可能(reconfigurable)多機能ALUパイプラインを含む多くのノードから構成されるコンピュータに関する。

#### 関連技術

本発明のコンピュータは、パラレルパイプライン式のコンピュータである。先行技術は、パラレルリズムとパイプラインニングという概念を成る限

られた文脈において開示している。例えば、米国特許第4,589,067号を参照されたい。

しかしながら、本発明の内部アーキテクチャは、全部ではないとしても、殆どすべてのコンピュータ構築ブロック(building block)が同時にアクティブであることを許容するという点において特徴を有している。

米国特許第4,589,067号は、動的(dynamic)に再構成可能なALUパイプラインに基づくベクトルプロセッサについて述べている点において先行技術の典型的なものである。このプロセッサは、本発明の再構成可能パイプラインの単一の機能ユニットに似ている。或る意味において、本発明のノードのパイプラインは複数のパイプラインのうちの1つのパイプラインである。本発明と比較するのに恐らく値するその他の構造には、Kungのシストリックアレイ(Systolic Array)概念、MITのデータフロー概念、及びその他のパラレルアーキテクチャがある。

カーネギーメロン大学のH. T. Kungによる

シストリックアレイ概念は、コンピュータを“波”のように通るデータを含んでいる。本発明とは異なり、シストリックアレイシステムは、各構築ブロックが与えられた演算を実行する同種の構築ブロックから成っている。シストリックアレイコンピュータにおいては、データが流れると、同一の構築ブロックの間の連結は、演算の間、固定されたままである。せいぜい、構成(configuration)は総てのデータがシストリックアレイによって処理されるまで変えることができないというだけである。一方、本発明においては、構築ブロック間の接続は何時でも変えることができ、データがパイプラインを通り抜けるときでもこれができる(すなわち、ダイナミック連結の再配列)。本発明は、また本発明のノードパイプラインの各構築ブロック(すなわち、機能ユニット)が、隣りのものとは異なる作動を行うことができるという点においてシストリックアレイ概念とは区別される(例えば、機能ユニット1—浮動小数点式掛算; 機能ユニット2—整数引算; 機能ユニット3—論理的比

較等)。

さらに、演算の進行中、本発明の各構築ブロックは異なる機能を行うことができる。

MITデータフローコンピュータは、パイプラインの配列内で接続できるハードウェア呼び出し(invoked)のネットワークで構成されている。命令処理は“データフロー”と非同期である。各データワードは、データの適当な命令ユニット(instruction unit)への経路指定を決定するトークンビット(token bit)のフィールドが付加されている。各命令ユニットは、各オペランド入力(operand input)のデータ列を有している。この命令は、総てのオペランドが存在するに至るまで“点火”(すなわち、実行)されない。本発明は、データの演算の実行をする(例えば、データを処理する命令として作用する)ハードウェア機能ユニットのパイプラインネットワークを流れるデータの概念を含んでいる。しかしながら、本発明は非同期モードにおいては働かない。その代わり、データはメモリーから持って来られ超高速マイク

ロシーケンスユニットの中央に集められる制御装置を通してパイプライン化された命令ユニットにスイッチ(MASNET)によって経路指定される。この同期制御シーケンスは、データフローアーキテクチャによって呼び起される非同期的分配データルーティング(asynchronous distributed data routing)と鮮やかな対比をなす。

さらに、本発明は、データフローマシン(Data Flow Machine)と異なり、トークンフィールド(すなわち、適切な機能ユニットにデータを案内するデータフィールド)を有していないし、また機能ユニットは複数の列(すなわち、複数のオペランド、命令又は結果を保持するバッファ)を有していない。データフローマシンは、データを待機する機能ユニットを有している。本発明は、連続的にアクティブである機能ユニットを有している。本発明のパイプラインの制御は、マイクロシーケンサと称するセントラルコントローラによって達成され、一方、データフローマシンは分配された制御装置を使用する。本発明は、又、TA

Gフィールドを使用してデータの内部的な流れに基づいてそれ自身を再構成する能力を有し、これはデータフローマシンには見い出せない特徴である。

さらに、データフローコンピュータは連続的に流れるベクトルデータに対する一連の類似又は否類似の演算を効果的に達成できない(すなわち、パイプラインを流れる総てのデータ上の単一の演算機能しかない)。これに対し、本発明はこの計算を極めて自然に達成できる。

本発明のバラレルアーキテクチャと、他のバラレルアーキテクチャとの間には他に2つの基本的相違点がある。第1に、本発明の各ノードは、独特のメモリー/プロセッサ構想(構成)を含んでいる。他のバラレルアーキテクチャは、隣りのノードと連結するために増設される既存のスタンドアロン(stand-alone)コンピュータアーキテクチャを含んでいる。第2に、他の通常のマルチプロセッサ/バラレルコンピュータは、全体的(global)な通信(communication)の間、局部

的(local)な処理が一時中止されるようにプロセッサ間の伝達を監視・制御するための中央演算処理装置(CPU)を使用する。本発明のノードは、インタープロセッサルーター(interprocessor router)と、データの局部的な処理を妨害しないで通信を許容するキャッシュメモリー(cache memory)を使用する。

以下の米国特許は、プログラマブル或いは再構成可能パイプラインプロセッサについて論じている。: 3, 787, 673; 3, 875, 391; 3, 990, 732; 3, 978, 452; 4, 161, 036; 4, 161, 036; 4, 225, 920; 4, 228, 497; 4, 307, 447; 4, 454, 489; 4, 467, 409 と 4, 492, 953。プログラマブルとノンプログラマブルの両方のパイプラインプロセッサの歴史について有用な議論が米国特許4, 598, 655号のコラム1乃至コラム4に見出される。さらに、マイクロパイプラインコンピュータの初期の開発努力についての適切な議論が、Peter M. Kogge による「パイプラインプロセッサのプログラミング」と題する、

1977年3月号の「コンピュータアーキテクチャ」の63頁乃至69頁の記事に見出される。

最後に、次の米国特許をパイプライン化されたプロセッサについての一般的な議論として引用する：4,051,551；4,101,960；4,174,514；4,244,019；4,270,181；4,363,094；4,438,494；4,442,498；4,454,578；4,491,020；4,498,134と4,507,728。

#### 発 明 の 要 約

簡潔に説明すると、本発明は、並行動作する強力なノードを少数（例えば128）使用するものである。個々のノードは、その必要はないけれども、同時性を持たせる(synchronized)ことはできる。ノードの数を制限することによって、全体的な通信及び如何なる与えられた問題を解決するためにも要求とされる関連するハードウェアとソフトウェアの関係は驚き易いレベルに保たれ、同時に並行性に固有のものであるゲイン(gain)とスピード(speed)と容量(capacity)を有利に使用できる。さらに、本発明のノード間のインタープロ

セッサ通信は実際起こるのであるが、ノード内のデータの局所的な処理を速くすることはない。これらの特徴が、多量のデータの迅速で大変効率のよい処理手段を提供する。本発明の各ノードは、速さと性能においてクラス6のスーパーコンピュータ（例えば、Cray 2 Cyber 205等）に匹敵する。与えられたノードにおいて、コンピュータは与えられたノードの各クロックサイクル中に、仮にすべてでないにしても機能ユニットの殆どのものがアクティブであるように同期的な、動的に再構成可能なパイプラインで組織された数多くの（例えば30の）機能ユニット（例えば、浮動小数点算術プロセッサ、整数算術／論理プロセッサ、特定用途プロセッサ等）を使用する。このアーキテクチャのデザインは、メモリー内の中間的結果の記憶量を最小にするのに役立つとともに、典型的な計算の持続するスピードが機械の最大のスピードに近くなるように保証している。これは、例えば、与えられた演算に対する現実の持続するスピードが機械の最大のスピードよりはるかに遅い

現行のクラス6のスーパーコンピュータの場合にはないことである。さらに、本発明は、マルチプルメモリープレーン、動的再構成可能なパイプライン及びインタープロセッサデータルーター間のフレキシブルで一般的な連結関係を提供する。

本発明の各ノードは、算術／論理ユニット(ALU)、マルチプレーンメモリー(multiplane memory)及びメモリープレーンと再構成可能なALU間のデータのルーティングのためのメモリー-ALUネットワーク(MASNET)を含んでいる。各ノードはまた、各ノード内に演算のタイミングと性質を指示するためのマイクロシーケンサとマイクロコントローラを有している。ノード間の通信は、複数のハイバースペースルーター(hyperspace router)によって制御される。重要なオフライン大容量記憶装置(off-line mass storage)と関連づけられている前置(front end)コンピュータが、マルチノードコンピュータに入力命令(instructions)を与える。ノードの好適なトポロジー接続は、ブール(boolean)ハイパーキュー

ブのそれである。

各ノード内の再構成可能なALUパイプラインは、浮動小数点プロセッサ、整数／論理プロセッサと特定用途要素(special-purpose element)を有している。プロセッシング要素は、多くの使用例が知られているサブストラクチャに結線される。3つのハードワイヤードサブストラクチャ(hard-wired substructure)が、再構成可能なALUパイプライン内にしばしば現れる。1つのサブストラクチャは、1つの2要素式ユニットから成り、もう1つは1つの3要素式ユニットから成り、最後のサブストラクチャは1要素式ユニットから成っている。3要素式サブストラクチャは、2要素式サブストラクチャの2倍の頻度で通常見出され、2要素式サブストラクチャは、1要素式サブストラクチャの2倍の頻度で見出される。これらのサブストラクチャを有効に使用すると、ALUパイプラインの構成をコントロールするために使用されるスイッチングネットワークの複雑さを少なくすることに役立つ。

本発明は図面を参照することによってより理解されるであろう。

#### 実 施 例

この詳細な説明中、同じ符号は、本発明を説明する違った図面でも同じ要素を表わすために使用される。

第1図で示される本発明の好ましい実施例によるコンピュータ10は、ノード12として言及される多くのマルチプルメモリー/演算ユニットを含んでいる。コンピュータ10は、高ベクトルとスカラの、効率(efficiency)及びスピードを備えた算術及び論理演算を実行可能な並行処理のできる汎用性を備えている。このような装置は、広範囲の演算問題を解くことができる。夫々のノード12は、種々のユーザーのプログラムの開発、マルチノードの操作及び演算、オフラインデータの処理に適した高い汎用性を備えた前置コンピュータ16にドロップラインネットワーク18を介して接続されている。前置コンピュータ16は、インターコネクション22によってオフライ

ン大容量記憶ユニット20に接続されている。夫々のノード12は、また、ノード間コネクション14によって隣りのノードに接続されている。理解と説明のために、25基のノード12だけが、第1図に簡単なノード間リンク14とともに示されている。しかしながら、ノード12は、一般的なハイパーキューブ構成によって接続されてもよいし、本発明は、適用例に応じて要求される128よりも多い又は少ないノードからなっていてよいことがわかるであろう。他の従来例の並行処理コンピュータに見られるような、多数の比較的低速のマイクロプロセッサを相互接続するのではなく、本発明は、比較的少数の、相互接続された大容量、高スピードの強力なノード12を用いるものである。本発明の好適な実施例によると、その構成は、典型的には、1乃至128個のノードからなっている。このアプローチは、複数のノード12の間の物理的な、及び、論理的なインターコネクション14の数を限定する。好ましい接続のトポロジーは、ブールのハイパーキューブ(hypercube)の

接続である。コンピュータ10の夫々のノード12は、処理スピードと処理能力においてクラス6のスーパーコンピュータに匹敵する。

典型的な個々のノード12の詳細は、第2図に示されている。夫々のノード12(これは、コンピュータ10の構築ブロックであるが)は、5つの基本要素からなっている。即ち、(1)多くの(例えば9からそれ以上の)高性能及び特定用途要素62を有する再構成可能なALUのバイブライン24、(2)独立メモリープレーン30のグループ28、(3)ノンブロッキングマルチプル入力とマルチプル出力のスイッチMASNET(メモリー/ALUスイッチ・ネットワーク)26、(4)マイクロシーケンサ40及び(5)マイクロコントローラ42である。第2図は、メモリーALUネットワークスイッチ(MASNET)26によって再構成可能なバイブライン24に接続する8個のメモリープレーン30からなるノード12をして示している。本明細書では、“プロセッシング・ユニット”、“機能ユニット”、“プログラマブルプロセッサ”及び“構築ブロック”

”という用語は、浮動小数点算術プロセッサ、整数/算術/論理プロセッサ、特定用途プロセッサ、又はこれらの組合せからなる算術/論理ユニット62を意味する。

マイクロシーケンサ40は、メモリー28、MASNET26と再構成可能なALUバイブライン24の夫々にライン46を介して接続されている。同様にマイクロコントローラ42は、同じ要素にライン44を介して接続されている。マイクロシーケンサ40は、種々の要素の間又はその中でデータのクロッキングを支配し、ノード12の夫々のクロックチック(tick)のためのデータの経路(pathways)とバイブライン24の構成(configuration)を定める役割を果たしている。典型的な操作では、オペランドの新しいセットがバイブライン24に供給され、新しい結果のセットがノード12のあらゆるクロックにおいてバイブライン24から得られる。マイクロシーケンサ40は、バイブライン24、MASNET26、メモリープレーン30の構成を定めるマイクロコードを選択する責任



を負っている。典型的な操作では、アドレスは、特定の終端アドレスに到達するまで、特定の出発アドレスから夫々のクロック時間中に連続して増加する。アドレスランプ (ramp) は演算終端の割り込みフラッグが出されるまで連続して繰返される。メモリー28の与えられたプレーン30によって使用される実際のメモリーアドレスは、選択されたアドレスモード40に依存してマイクロシーケンサ40のアドレスとは異なっている (メモリープレーンについては後述を参照)。

ノード・マネジャーとも呼ばれるマイクロコントローラ42は、ノード12の各部分の初期化とベリフィケーション (verification) のために使用される。与えられた演算のために、最初のセットアップ後、コントロールはマイクロシーケンサ40に送られ、これが演算が完了するまで引き継ぐ。原理的には、マイクロコントローラ42は、演算が実行されている間作動している必要はない。しかし、典型的な操作においては、マイクロコントローラ42は演算の進行を監視したり、次の演

算のためのコンピュータの使用されていない部分を準備させているであろう。

最小の単位のノード12を構成する5つの基本要素に加えて、夫々のノード12は、ローカルな大容量記憶ユニット、グラフィックプロセッサ、前処理又は後処理プロセッサ、副データルーター、及びそれらの類似物を含むように拡張されてもよい。夫々のノード12は、ノードマネジャー42がスタンドアロンタイプのマイクロコンピュータであるため、スタンドアロンタイプのノードとして操作可能である。しかしながら、通常の場合、ノード12は前置コンピュータ16からプログラムされる。

1つのメモリープレーン30の配置は、第3図に模式的に表されている。メモリープレーン30は高い容量を有し、装置10のクロックにおいてデータワードを取り出し (読取) 又は送り (書込) することができる。各メモリープレーン30は、読取のみ、書込のみ、或いは読取/書込操作ができるようになっている。メモリープレーン30は3

つの可能なアドレスモード、すなわち、(1)直接型 (direct)、(2)翻訳型 (translate)、(3)演算型 (computed) のモードを有している。すべての3つのモードについて、ワーキングアドレスは、コンピュータ10の前のサイクルにおいて、先取り (prefetch) アドレスレジスターによって先取りされる。直接型モードでは、マイクロシーケンサのアドレスバス46からのアドレスが、対象となるメモリー要素を選択するために使用される。翻訳型モードでは、マイクロシーケンサアドレスは、アドレスの大きなメモリーテーブル中の実行中のアドレスを見出すために使用される。このアドレスの大きなテーブルは、翻訳メモリーバンク即ちテーブル50として示されている分離型メモリーユニット中に記憶される。翻訳テーブル50は、メインメモリーバンク54を通じ任意の走査パターンを発生させるために使用することができる。これはまた、或る指定されたメモリー要素が書き換えられることがないように守るためにも使用できる。演算型アドレスモードは、パイプライン24

が次に読取或いは書込されるデータワードのアドレスを定めることを許容する。

再構成可能パイプライン24は、第4図にユニット62として示される種々のプロセッシング要素と、第5A図、第5B図にフローネット (FLONET) 70として示されるスイッチ網からなっている (FLONETは、Functional and Logical Organization Network の略語である)。3つの永続的ハードワイヤ型 (permanently hardwired) サブストラクチャすなわちユニット62、64又は66がFLONETに接続されている。FLONET70は、第5A図中に68、そして第5B図中に69で集合的に示されているパイプラインサブストラクチャ62、64、66の結線を再構成 (reconfigure) する。特定の再構成可能な相互接続は、新しい再構成がノード12のクロック時間内に形成できるように電子スイッチによって達成される。特定の状況におけるハイレベルのデータ処理の例が第4図に示されている。パイプライン型プロセッシング要素は、浮動小数点算術プロセッサ (例えばAMD29325、

Weitek 1032/1033)、整数算術/論理ユニット62(例えばAMD29332)、ベクトル再生ユニット或いは収束検査機(convergence checkers)のような特定用途要素からなる。前述の特定用途要素に関する有用な議論がD. M. Nonsechuck, M. G. Littman, W. FlanneryによるJ. Sci. ComputeのVol.1No.1(1986)の“ナビエーストークスのコンピュータのミニノード上における二次元の非定常粘性流体のシュミレーション”と題する記事の中に見出される。プロセッシング要素62は、多くのユーザーの応用プログラムにしばしば使用されている3つの個別のサブストラクチャ62、64、66とともに結線されている。最も一般的に使用されるサブストラクチャ64、66の二つは、第4図に点線で囲まれた要素として示されている。サブストラクチャ64は、4つの入力と1つの出力を有する3つのALUユニット62からなっている。2つのALUユニット62は2対で合計4つの入力を受ける。2つのALUユニット62の出力が、第3のAL

Uユニット62の2つの入力を形成する。3つのALUユニット62の夫々は、浮動小数点及び整数付加、減算、乗算、及び除算、論理的AND、NOT、OR、NOT、及び排他的OR、マスク、桁送りを行うことができ、定数を記憶させるために使われる論理的レジスタファイルとともに比較機能を行うことができる。サブストラクチャ66は、2つの算術/論理ユニット62からなっており、3つの入力と1つの出力を受けもつようになっている。2つの算術/論理ユニット62の1つは、2つの入力を受け、第2の算術-論理ユニット62のための1つの入力を形成する出力を作り出す。第2の算術/論理ユニット62に他の入力が直接外側から送りこまれる。サブストラクチャ66の1つの入力は第2の算術/論理ユニット62から供給される。したがって、サブストラクチャ62は、3入力、1出力の装置からなっている。第3のそして、最後の最も一般的なサブストラクチャは、一個のスタンドアロン算術/論理ユニット62であり、2つの入力と1つの出力を有している。サ

ブストラクチャ62、64、66は、これら夫々の構成に永続的に結線されている。しかしながら、これらのユニットの再構成はFLONET70によって制御される。簡略化されたFLONET70が、模式的に第5A図に示されている。簡略化のため、2つの3要素式サブストラクチャ64、2つの2要素式サブストラクチャ66及び2つの1要素式サブストラクチャ62が図示されている。この構成によって、12の機能ユニットのハイレベルな再構成可能パイプライン24を得ることができる。

第5B図は、FLONET/ALUの接続の最適状態のレイアウトを示している。本発明の好ましい実施例10によると、3要素式サブストラクチャ64と2要素式サブストラクチャ66の最適比は、1.5~2.0:1の範囲にある。同様に、2要素式サブストラクチャ66と1要素式サブストラクチャ62の最適比はおおよそ2:1である。したがって、第5B図は、8つの3要素式サブストラクチャ64と4つの2要素式サブストラクチャ66と2つの1要素式サブストラクチャ62からなる最適な

図を示していることになる。3要素式サブストラクチャ64の数は第5B図で示される実施例では、6~8の間で変更することができる。前述した好ましい比はおおよそその値であり、実際の使用に際してわずかに変更してもよい。しかしながら、前述の比が、実際に最適に近い結果を示すことがわかった。

本発明の好ましい実施例によると、第5B図のサブストラクチャ62、64、66の集合69は後述のように組織される機能ユニット即ち、構築ブロック62を有している。すなわち、8つのサブストラクチャ64中の3つの機能ユニット62(即ち、プログラマブルプロセッサ)の夫々は、AMD29325のような浮動小数点プロセッサであり、サブストラクチャ66の2つは、AMD29325のような浮動小数点プロセッサの形式の2つの機能ユニット62を夫々有しており、一方、残りの2つのサブストラクチャ66は、AMD29332のような整数/論理プロセッサを有しており、最後に、残りの機能ユニット62の1

つはAMD29325のような浮動小数点プロセッサであり、他の1つの機能ユニット62はAMD29332のような整数/論理プロセッサである。或いは、ハイブリット機能ユニット62を形成するようにプロセッサを対にすることもできる。例えば、AMD29325のような浮動小数点プロセッサは、同業者によって周知の方法によって、AMD29332のような整数/論理プロセッサと一対にすることができる。これによって、機能ユニット62は、浮動小数点型と整数/論理型間の切り換えが可能となる。又、同様の結果を得るために、Weitek3332のような1つの多機能プロセッサ(浮動小数点型、整数算術/論理型)を使用することもできる。

MASNET26(すなわち、Memory Alu Switch Network)は、第6図で、16の入力と16の出力を有するものとして詳細に示されている。MASNET26は、如何なる入力の出力に対する接続もノンブロッッキングであるようにビーンズ(Benes)スイッチネットワーク配列に交差して接続されバイブラ

ためにも使用される。このルーティング(routing)は、付加的な遅延を導入することなくMASNET26を通じてデータが流れる時に達成される。同様に、与えられたノードのハイバースペースルーター80は、遅延の導入なしに、必要な時にデータストリーム中に境界点値(boundary point value)を組み入れることができる。ノード間通信の詳細な説明は以下の通りである。

マルチノードコンピュータ10の全体的なトポロジーは、ハイパーキューブのトポロジーである。ハイパーキューブは、任意のノード間通信のために必要とされる時間とノード12の間の物理的な接続の数の間の妥協を表している。2つのアドレスモードがノード間データ通信、即ち、(1)全体的(global)なアドレス及び(2)明示(explicit)のパウンダリ・ポイント・デフェニションすなわちBPDをサポートする。全体的なアドレスは、単純に拡張したアドレスであり、ここでアドレスがデータのノード/メモリープレーン/オフセットを特定づける。ソフトウェアの見地からは、アドレス

イン化された、レジスターファイル72(例えば、Weitek1066)から構成されている。第6図に示されるMASNET26は、 $16 \times 16$ の回路である。夫々のレジスターファイル72がローカルメモリーを有していることは、また、MASNET26を使用することによりネットワークを通じて流れるデータに再指令を与えることができることを意味している。この特徴は、例えば、共通のソースから2つのデータの流れを作り出し、種々の要業によって一方が他方に対して遅延させられるようにするためにも用いることができる。共通のソースから複数データの流れを形成できることは、MASNET26のもう1つの特徴である。第7図は、 $2 \times 2$ のMASNET(即ち、1つのレジスターファイル72)が如何にしてこれらの簡単な2つのタスクを実行するかをより明確に示している。

MASNET26は、また、ハイバースペースルーター80を通じ、境界の(bordering)ノード12にノードの境界線に対応するデータワードを経路指定する(route)ノード間通信(communication)の

はコンピュータ10中でそのレンジ全てのノードに(across)及び線形の単純なアドレスとして取扱われる。ノード間通信はソフトウェアによって処理され、デフォルト・アービトレーション及びコミュニケーション・ロック・パラメータが選択された場合、プログラマーには明確にわかる。BPDは、バウンダリ・ポイントの明示のデフェニション、それらのソース及びすべての目的地アドレスを含んでいる。BPDデータが発せられたときはいつでも、直ちに第8図に示される目的地ノード12中のBPDキャッシュ82に経路指定される。ローカルアドレスとBPDは混成されてもよい。BPDは、他のノードによってデータが要求される前に予めバウンダリ・ポイント・データを通信することによって、ノード間通信のオーバーヘッドを殆ど除く能力をもっているが、BPDに優る全体的アドレスの主な利点は、ソフトウェアを簡易化することである。

データは、夫々のノード12に取付けられているローカルスイッチネットワークを使用してノー

ド12の間を物理的に経路指定される。ハイバースペースルーター80として前に言及したローカルスイッチネットワークが、第8図に表されている。ハイバースペースルーター80は、ビーンズネットワークに似たトポロジーの、ノンブロッキング・パーミューテーション・ネットワークである。マルチノードクラスのコピュータすなわち等級 $d$  (即ち、 $NN=2^d$ 、 $NN$ はノードの数)のためには、ハイバースペースルーターは $d$ 個の隣り合うノード12に、ホストノード12のための1つの付加的な入力をプラスし、 $d+1$ の入力を可能としている。データは、データが持っている目的地アドレスがハイバースペースルーターのスイッチ状態を設定するために使用される点においてセルフルーティング(self-routing)である。8ノード式システムが第8図に示されている。この例では、 $d=3$ であり、夫々のハイバースペースルーター80は3つのマイナークロック遅延を有する、 $4 \times 4$ のネットワークを有している。 $d$ が整数であり、 $3 < d < 8$ の場合では、 $8 \times 8$ のルー

ースペースルーターは $16 \times 16$ のスイッチに増える。検索テーブルは限界を越えて大きくなるから、パーミューテーション・ルーティングは、検索テーブルよりもいくらか遅いビットスライス(bit-slice)ハードウェアによって達せられる。これらの考え方に基づき、当初の好ましいコンピュータ構成として128のノードを設定した。

ノード12間のデータ伝達は、1ギガバイト/秒のduplex比でバイト・シリアル・フォーマット中の光ファイバケーブルで行われる。この速さは時々起こるバースト伝達と、また、将来のコンピュータの拡張のための略々2倍規模のヘッドルームを与える。夫々のノード12は、キャッシュバスサイクルのためのホストノード要求がない場合でも、ハイバースペースルーター80により連続的にデータが増加する1つのMワード(Mword)バウンダリ・ポイントとライト・スルー(write-through)キャッシュを有している。かくして、現在のバウンタリ・データはALUパイプライン入力に物理的に論理的に近く維持される。

ター80が必要であり、 $d=7$ であると完全なスイッチ利用ができる。ハイバースペースルーター8は $2^8=1$   $d=1$ の出力のために構成されなければならないので、最適なハードウェア性能は、次のサイズを有するコンピュータアレイによって与えられる。

$$NN = \frac{2^{2n}}{2}, \quad n = 0, 1, 2, 3 \dots$$

1, 2, 8, 128, ...個のノードの構成は、完全にハイパーキューブルーター80を利用する。非整数 $2^d$ を有するマルチノードコンピュータの構成も、また、ハイバースペースルーター80が次の整数の次元にスケールアップされる場合を除いてサポートされる。このことは、付加的なスイッチのハードウェアが負担となる点は別として、僅かに大きい量の記憶装置がパーミューテーションテーブルのために要求されるだけであるという点でそれほどシビアではない。ノードは、これらのテーブルを高速検索テーブル中に記憶する。テーブルの長さは $(d+1)$ である。コンピュータが128のノードを超えて、大きくなると、ハイバ

本発明は、好ましい実施例を参照して説明したが、本発明の趣旨及び範囲を外れることなく種々の変更が本発明を構成するある部分や方法についてなしうる事が理解されるであろう。

#### 4. 図面の簡単な説明

第1図は、ブールハイパーキューブ(boolean hypercube)の部分集合である、隣同志が2次元に配列されたグリッドの形に配列されたマルチノードコンピュータの実施例の図、

第2図は、メモリー/MASHENT/ALUの回路接続を図示した単一ノードの模式図、

第3図は、第2図に示すような単一ノード内の1つのメモリープレーンのレイアウトを描いた模式線図、

第4図は、各ノードの再構成可能なALUパイプ内に見出される5つの算術/論理ユニットで形成された2つの代表的なサブストラクチャ(sub-structures)の図、

第5図Aは、サブストラクチャの構成を変えることを許容する典型的なALUパイプライン機構

とスイッチングネットワーク (FLONET) の図、

第5図Bは、再構成可能ALUパイプラインにおいて3つの普通のサブストラクチャのグループに対するFLONETの接続の好適実施例の図、

第6図は、ブロックが6ポートレジスタファイルを表わす、3.2-レジスタ $\times$ nビットの、メモリ/ALUネットワークスイッチと、ノード間通信ユニットの概略線図、

第7図は、“p”要素の相対移動により如何にして入力データが2つの出力データを取り出せるかを図示した2 $\times$ 2 MASNET の模式図、

第8図は、ハイバースペースルーターの各ノードの MASNET ユニットに対する関係を示した8-ノードハイバークューブの模式図である、

- 10...コンピュータ
- 12...ノード
- 16...前置コンピュータ
- 18...ドロップラインネットワーク
- 20...オフライン大容量記憶ユニット
- 24...再構成可能ALUパイプライン
- 26...MASNET

- 30...独立メモリープレーン
- 40...マイクロシーケンサ
- 42...マイクロコントローラ
- 50...翻訳テーブル
- 54...メインメモリーバンク
- 70...FLONET
- 72...レジスタファイル
- 88...ハイバークューブルーター

FIG. 1

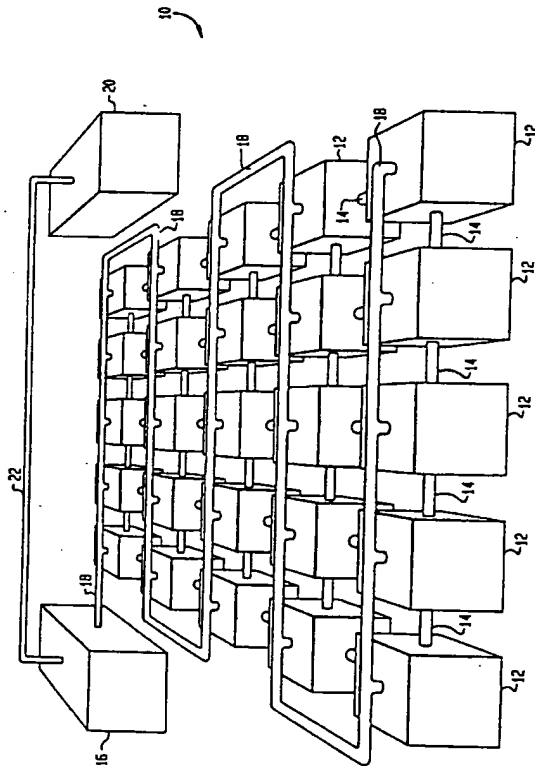


FIG. 4

オペランド (入力)

